(1)

Research Note 81-28

DEVELOPMENT OF A USER'S GUIDEBOOK FOR TRAINVICE II

Robert W. Swezy
Science Applications, Inc.

and

Robert A. Evans
Army Research Institute

SIMULATION SYSTEMS TECHNICAL AREA

U. S. Army
Research Institute for the Behavioral and Social Sciences

December 1981

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Research Note 81-28	2. GOVT ACCESSION NO. AD-A128 143	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DEVELOPMENT OF A USER'S GUIDEBOOK FOR TRAINVICE II		5. TYPE OF REPORT & PERIOD COVERED Final Report
		6. PERFORMING ORG. REPORT NUMBER SAI-80-066
7. AUTHOR(s) Robert W. Swezey Robert A. Evans		8. CONTRACT OR GRANT NUMBER(s) MDA 903-79-C-0428
9. PERFORMING ORGANIZATION NAME AND ADDRESS Science Applications, Inc. 1710 Goodridge Drive McLean, Virginia 22102		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q162743A772
11. CONTROLLING OFFICE NAME AND ADDRESS Simulation Systems Technical Area - PERI-IE US Army Research Institute for the Behavioral and Social Sciences-5001 Eisenhower Ave. Alex, VA 22333		12. REPORT DATE December 1981
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Monitored Technically by: G. Gary Boycan and Marshall A. Narva

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

TRAINVICE
TRAINVICE II
Transfer of Training Model
Training Device/Simulation Assessment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report documents the effort to develop a User's Guidebook for the Application of the transfer of training model known as TRAINVICE II (See Research Note 81-29). In addition to presenting the development process and providing an overview of the guidebook, this report discusses a variety of technical concerns regarding the TRAINVICE II model. Recommendations are also provided in terms of future research and development for such approaches.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Research Note 81-28

DEVELOPMENT OF A USER'S GUIDEBOOK FOR TRAINVICE II

Robert W. Swezy
Science Applications, Inc.

and

Robert A. Evans
Army Research Institute

Submitted by:
Frank J. Harris, Chief
SIMULATION SYSTEMS TECHNICAL AREA

Approved by:
Joseph Ward, Acting Director
TRAINING RESEARCH LABORATORY

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

December 1981

Army Project Number
2Q162743A771

Training and Utilization
in Military Systems

PREFACE

Overview

This report documents an effort to develop a User's Guidebook for the application of the transfer of training model known as TRAINVICE II (Swezey and Evans, 1980). The Guidebook was developed as an initial product in a series of efforts to provide researchers with a methodology for evaluating training devices. A TRAINVICE-like approach may ultimately be designed to aid users in determining how well a training device covers and implements specific training requirements.

Organization of this Report

This Report is organized in the following sections:

- Preface
- Introduction
- Background
- Results
- Recommendations

Preface

This section briefly provides an overview of the effort, discusses the organization of this report and acknowledges contributions to the project.

Introduction

The introduction provides a brief description of the project. Required tasks and deliverables are also discussed.

Background

This section provides a historical discussion regarding the need for TRAINVICE II, accompanied by a brief description of several related models. A discussion of TRAINVICE II is also provided.

Results:

Various positive and negative aspects of the TRAINVICE II model are discussed in the Results section, and special features of the Guidebook are also mentioned and discussed.

Recommendations:

The final section summarizes findings and technical issues concerning the TRAINVICE II model and presents recommendations and considerations for possible further research.

Contributors:

Several individuals have made major contributions to this effort. Marshall A. Narva, of the U. S. Army Research Institute for the Behavioral and Social Sciences (ARI), developer of TRAINVICE II and the original Technical Monitor for this effort, provided major guidance to the project staff. G. Gary Boycan, Dr. Narva's successor as Technical Monitor, also contributed substantially to this project. Robert G. Wanschura of ARI has made a thorough and constructive critique of the draft Guidebook for which the authors are most grateful. Claude G. Songy, of Science Applications, Inc., was a major contributor to the early drafts. He further assisted in this project by revising the original TRAINVICE II index. It is Mr. Songy's revised index that is used in the Guidebook. Dennis G. Faust provided major support by reviewing and commenting on preliminary versions of the Guidebook manuscript.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

TRAINVICE II is a modification of an earlier training device effectiveness model known as TRAINVICE (Wheaton, et.al., 1976a, 1976b, and 1976c). TRAINVICE has been applied in various Army, and other training evaluation efforts (see for example, Swezey, et.al., 1977a, 1977b,; Klein, et.al., 1978). For a variety of reasons, ARI decided to revise the original TRAINVICE model, and the revision presently known as TRAINVICE II is the result of this decision.

The development of TRAINVICE II represents a major stride forward in the overall ARI device effectiveness assessment program. In order to facilitate use and further development of the model however, it was necessary to document TRAINVICE II in a format which would enable potential users to apply the model during various training equipment acquisition and/or evaluation exercises. The development of a user's Guidebook, therefore, was the primary focus of this effort.

The work conducted under this project involved four basic tasks:

- Task 1 - Develop a Draft User's Guidebook for TRAINVICE II
- Task 2 - Expert review of the Draft Guidebook
- Task 3 - Revision of the Draft Guidebook per expert comments
- Task 4 - Final Report

Task 1 - Develop a Draft User's Guidebook for TRAINVICE II

During this Task, all relevant TRAINVICE documentation was reviewed. A draft Guidebook for users was developed in eight Chapters, under close cooperation with the technical monitors. The chapter topics are briefly reviewed below.

Chapter 1 - Introduction: How To Use The Guidebook
In Chapter 1 a general introduction to the purpose of, and concepts used in, the TRAINVICE II model is provided. Special features of the Guidebook are called to the reader's attention.

This chapter also provides a review of the underlying rationale for the model and a brief discussion of its history, as well as a discussion of previous developmental efforts associated with TRAINVICE II. Chapter 1 describes the situations in which TRAINVICE II is applied and presents generalized guidance as to what, where, when, why and how to use the model.

The Chapter is written in narrative format and is presented in an easy-to-read fashion. Guidance on the development of the final draft was provided by ARI's Guidebook For The Development of Army Training Literature (Kern, et.al., 1977).

Chapter 2 - Coverage Requirements Analysis

In Chapter 2, as with the other component-oriented chapters, Narva's (1979a., b.) TRAINVICE II model is described in a procedural, how-to-do-it fashion. First, step-by-step procedures for conducting the analyses are described. This is followed by an example which assists the user in seeing how the steps in the analysis are applied in a specific hypothetical situation. Flow charts and completed worksheets are provided to aid users in applying the analytic procedures to actual training device evaluation settings.

Chapters 3 through 8 - Remaining TRAINVICE II Components

Each of the remaining chapters present, in a fashion similar to Chapter 2, procedures for applying the TRAINVICE II components.

The titles of these chapters are:

Chapter 3 - Coverage Analysis

Chapter 4 - Training Proficiency Analysis

Chapter 5 - Learning Difficulty Analysis

Chapter 6 - Physical Characteristics Analysis

Chapter 7 - Functional Characteristics Analysis

Chapter 8 - TRAINVICE II Computation

Appendices are also provided to assist users in calculating the TRAINVICE II Index. Appendix A presents Behavioral Category Definitions, adapted from TRADOC Pamphlet 350-30, which are used in the Physical and Functional Characteristics Analyses portions of TRAINVICE II. Appendix B provides learning Guidelines for Good Instructional Practice, which were extracted from TRADOC Pamphlet 350-30, and which are used in conjunction with the Behavioral Category Definitions. Appendix C provides a brief discussion of potential prescriptive applications of TRAINVICE II (uses of the model with a concept rather than with an existing device). A procedures checklist, for application of the model, is presented in Appendix D. This Appendix, intended to serve as a performance aid for the experienced TRAINVICE II user, provides a step-by-step checklist for completing each component of the TRAINVICE II model. Finally, blank worksheets for computation of various model components are provided in Appendix E.

Presentation of each model component in a separate Guidebook chapter allows a user to calculate separate components of the model as desired. This format was based on past situations where only certain components of the TRAINVICE II model were used in an evaluation project (Klein, et.al., 1978). Each chapter of the Guidebook, therefore, stands alone to allow for separate applications of model components.

Task 2 - Expert Review of the Draft Guidebook

During Task 2, the Draft Guidebook was submitted to a small group of experts for review. The reviewers were knowledgeable about and interested in TRAINVICE II. Reviewers were asked to read and comment on each chapter of the Draft Guidebook.

The major review focus involved the following areas:

- Readability
- Ease of Use
- Level of Detail
- Degree of Proceduralization
- Clarity of Presentation
- Theoretical Issues

Task 3 - Revision of Draft Guidebook per Expert Comments

During this task, the Draft Guidebook was revised to accommodate comments provided by the reviewers. Every attempt was made to consider all comments and to make appropriate changes.

Task 4 - Final Report

This Task required technical documentation and comment on the Guidebook and on theoretical considerations involving the TRAINVICE II model. This report satisfies the Task 4 requirement.

BACKGROUND

The overall mission of the modern Army requires a broad range of training alternatives. In satisfying these alternatives, the Army increasingly relies on training equipment for instructional purposes, rather than on operational hardware. Various reasons exist for this shift. Included among these are the effects of reduced budgets, resulting in less availability of operational equipment for training purposes. Such reasons, although themselves sufficient justification for the shift to training devices, represent a non-instructional rationale.

There are, however, additional justifications deriving from the field of instructional technology. Training devices may, in various situations, actually be designed to train individuals better than can be done with operational equipment. Hands on experience with real hardware is often not necessary in order to learn its operation or maintenance. Training devices can be specifically designed to provide such instructional benefits as immediate feedback, reinforcement for correct responses, hierarchical sequencing of learning objectives from easy to difficult, measurement of achievement, and other positive features in complex skill learning environments. Inclusion of such features represents an improved learning situation not commonly found in operational equipment. Operational equipment and training devices are designed for distinctly different purposes.

The so-called "systems approach to instruction", presently widely accepted, greatly increases the importance of training devices as instructional tools. Simulators and other training devices, more than operational hardware, particularly fit into a systems approach to training. Such devices can be designed to meet even rigorous instructional requirements. Training devices can, for example, be designed to present instructional materials which can accommodate to individual differences and learning styles. Certainly, such flexibility cannot typically be incorporated in operational equipment.

While a very strong case can be made for the use and contributions of training devices, such devices are not without liabilities. A major concern involves cost. Developing and producing training devices may actually exceed the cost of operational equipment they were designed to simulate.

A second major concern involves uncertainty relative to the effectiveness of training devices. Instructional technology has been somewhat limited in its examination of the concept of transfer of training. Other factors such as unknown relationships between learning and training media, also tend to confound the effectiveness issue. Optimal training systems cannot be designed until these, and other similar issues, are resolved.

Such issues create a paradox, where reliance on training devices continues to increase, but their effectiveness remains uncertain. Efforts continue to be undertaken to resolve this problem. Various models have been developed which address the issue of predicting effectiveness of training devices. Some of these attempts are briefly mentioned here. For a recent discussion of this topic, see Adams (1979).

TRAINVICE

A recent major effort which addressed this issue has been described by Wheaton, et.al., (1976), with the development of the original model which has come to be termed TRAINVICE. TRAINVICE was designed to provide a framework for estimating the effectiveness of training devices.

The model combines judgmental data concerning several major variables into a figure of merit, tau (τ) which assesses training device to operational equipment transfer on a scale ranging from 0 to 1. Input to the model consists of three general kinds of information: (1) a detailed statement of the training objective, (2) a task analysis of the operational task, including relevant controls and displays as well as lists of appropriate skills and knowledges for each subtask, and (3) a similar task analysis, at the same level of detail, for the training task on the device or simulator.

Each input utilizes several variables, for which values are derived and combined into a figure of merit. These values are related to:

- Training Objectives
- Device and Operational Task Descriptions
- Device and Operational Subtask Description
- Device and Operational Subtask Controls and Displays
- Device and Operational Subtask Control and Display Functions
- Training Theory and Empirical Data

The Wheaton et.al. model essentially consists of three parameters, one of which has several components. The parameters combine in a multiplicative format to derive a final figure of merit. The parameters are:

Transfer Potential of the Training Device. This parameter is designed to provide information on the extent to which a training device is potentially (as opposed to actually) effective in providing position transfer to operational equipment. Transfer potential is divided into two components: Task Communality and Equipment Similarity. The Task Communality component involves a dichotomous decision for each subtask, as to whether or not the device enables a trainee to practice skills required for performance of that subtask on the operational equipment.

The Equipment Similarity component is itself sub-divided into two segments; termed Physical Similarity and Functional Similarity. Physical Similarity between the device and the equipment is rated for each subtask, using a four point behaviorally anchored rating scale defined in terms of just noticeable difference (jnd) units. The greater the physical discrepancy between training device and operational equipment for each subtask, the lower the rating.

Functional Similarity analysis is an assessment of the extent to which a training device involves the same type of information requirements as that required by the operational equipment. The scale is defined in

information theoretic terms. In each subtask, the number of bits of information is determined for each stimulus and response situation by estimating the number of states which may be assumed by a control or display. A similar assessment is made for the training device. A four point rating scale is then applied for each subtask, which assesses the functional similarity of the devices and equipment using a $\log_2$ unit. Summary values are then obtained for both Physical and Functional Similarity indices on each subtask.

Learning Deficit Analysis. The second major parameter in the Wheaton, et.al., model is termed Learning Deficit Analysis. This analysis involves a procedure where each subtask is assessed in terms of: (1) a five point rating of the criterion level of performance required for successful completion of each subtask, (2) a five point rating of the presumed input repertoire of trainees upon entry into training on the subtask, and (3) a ranking of subtasks on the difficulty (in terms of training time) of training the necessary skills and knowledges. The output of this stage of analysis is a number for each subtask indicating the deficit (that is criterion performance minus input repertoire) possessed by a typical trainee, weighted by the relative difficulty (in terms of estimated training time on the operational equipment) of surmounting that deficit.

Training Technique Analysis: The third major component involves an analysis of the particular skills and knowledges to be trained and their relationship to known relevant training principles and techniques. In this procedure, subtasks are first assigned to taxonomic categories using a 13 category task taxonomy developed by the U. S. Naval Training Device Center (1972). Second, the analyst evaluates a specified list of training principles and techniques and applies them to the taxonomically defined subtasks. The relevant training principles are then categorized as to stimulus, response and feedback aspects and the training device is rated on a seven point scale for each category.

Figure of Merit. The final step in the Wheaton, et.al., model is the calculation of a figure of merit (tau) which quantifies perceived training transfer. The model combines the judgmental data obtained from the model parameters into a figure of merit whose scale properties are ordinal. The figure can, therefore, be used in making relative judgments among alternative training devices.

The basic formula is:

$$\hat{T} = \frac{\sum_{i=1}^{n} (C \times S \times T \times WLD)_i}{\sum_{i=1}^{n} WLD_i}$$

$\hat{T}$	=	predicted transfer of training
C	=	task communality estimate
T	=	estimate of instructional effectiveness of the device
S	=	device similarity estimate
		and
WLD	=	estimated learning deficit to be overcome

During May, 1977, ARI convened a Conference for the purpose of reviewing the TRAINVICE model and to consider suggestions for simplification and revision. Invited to the conference were a variety of individuals who were directly interested in the model. The attendees consisted of representatives of the sponsoring agency, the developer, and other concerned agencies including both military and civilian organizations. (See: Swezey and Narva, 1978).

The TRAINVICE conference considered a variety of issues relevant to the TRAINVICE model during its three day meeting. Certain issues were resolve by consensus of conference participants; others were not. Several of the issues considered and resolutions, where they occurred, are listed in the following paragraphs.

1. It was determined that the TRAINVICE model fairly represents the state-of-the-art in training device assessment.
2. It was recommended that an estimate of mission criticality be considered for inclusion in future versions of the TRAINVICE model.
3. There was major discussion on the topic of how the Training Techniques analysis aspect of TRAINVICE might be simplified for easier use.
4. It was suggested that tau should best be evaluated on a relative basis, i.e., as opposed to other tau values on comparable items of equipment, rather than on an absolute basis.
5. It was determined that the scale properties of the TRAINVICE figure of merit are unknown.
6. It was recommended that TRAINVICE be considered a component of a much broader training device evaluation methodology since it does not consider such variables as: cost, maintainability, reliability, etc.
7. It was suggested that TRAINVICE might be best applied during the prototype design stage of the training equipment development cycle.
8. It was suggested that a method for assessing user acceptability of a training device be developed and included in the TRAINVICE model.
9. It was determined that TRAINVICE is as yet unvalidated as a methodology and that efforts in this area are necessary before TRAINVICE will achieve widespread acceptability or use.
10. It was pointed out that the output of a TRAINVICE analysis is extremely sensitive to the type of task analytic methodology used as an input. Ways of dealing with this problem should be developed.
11. It was suggested that efforts be devoted to develop ways of increasing the reliability and construct validity of TRAINVICE, as well as to ways of extending the model for application to more complex training situations.

TRAINVICE II

TRAINVICE II was generated as a revision of the original TRAINVICE model, (Wheaton et.al., 1976). The revised model was based partially upon experience gained in the use of TRAINVICE and partially upon the outcome of the TRAINVICE Conference. The primary goals of the revision were to increase the practicality and flexibility of the model's application and to adjust the model for a wider potential spectrum of users.

TRAINVICE II provides a means for answering three questions regarding training device use: "what", "why", and "how". These questions focus on units of behavior being trained by a device.

The "what" question refers to what should be represented in a training device. This issue requires two judgments. The first deals with the requirement for an activity to be represented in a device. The second helps determine if that activity is actually covered by the device. The purpose of these judgments is to determine if the training activities covered by a device actually address the requirements.

The answer to the "why" question attempts to provide the reasons for including training activities on a device. There are two aspects at this stage: Training Criticality, originally conceived as the degree of proficiency required at the end of training; and Training Difficulty, or the degree of difficulty expected for a trainee to reach the required level of proficiency. These factors, training critically and training difficulty, quantitatively weight the activities covered by a device.

The "how" question is related to the actual characteristics of the training device. That is, how well do the physical characteristics (i.e., displays and controls) follow specified quidelines for good instructional practice, and do the displays and controls adequately meet the perceived requirements of the device developer. The "how" question is then focused on the functional characteristics (capability of displays and controls) of the training device. That is, to what extent do the functional characteristics meet the guidelines for good instructional practice.

The TRAINVICE II components which correspond to the questions just reviewed are shown in Figure 1, below:

What?

Coverage Requirements Analysis

Should this unit of activity be represented in the device?

Coverage Analysis

Is this unit of activity represented in the device?

Why?

Training Criticality Analysis

What level of proficiency is required for this unit of activity?

Training Difficulty Analysis

How difficult will it be to train for this unit of activity?

How?

Physical Characteristics Analysis

How well do the Physical Characteristics of the device involved in training for this unit of activity meet the perceived requirements and applicable guidelines?

Functional Characteristics Analysis

How well do the Functional Characteristics of the device involved in training for this unit of activity meet the applicable guidelines?

Figure 1.

Judgments Made in TRAINVICE II (Narva, 1979b)

Each of these judgments is made for every skill or knowledge required in a training device with the values being substituted in the following formula (developed by Narva, 1979a, 1979b):

$$\frac{\Sigma\ (CR \times C \times Ci \times D) \times (PC + FC)\ i}{\Sigma\ (CR \times C \times 4 \times 4) \times (PC_{max} + FC_{max})\ i}$$

<u>WHERE</u>:

CR = Coverage Requirements Analysis Score

C = Coverage Analysis Score

C_i = Training Criticality Analysis Score

D = Training Difficulty Analysis Score

PC = Physical Characteristics Analysis Score

FC = Functional Characteristics Analysis Score

PC_{max} = Maximum Possible Physical Characteristics Score

FC_{max} = Maximum Possible Functional Characteristics Score

This equation yields an index ranging from 0 to 1, such that a larger index results in a larger number of required skills represented and higher ratings on the PC and FC analyses. This, in turn, presumable yields a greater potential for transfer of training from the device to the operational equipment.

Other Training Device Evaluation Models

Training Device Requirements Documents Guide

The Training Device Requirements Documents Guide (1979), presents a method for defining training requirements and for performing media analyses. The Guide provides decision rules to determine when a training device is required or when alternative media are more appropriate. When a device is determined to be appropriate, the Guide specifies how initial training device concepts are formalized, validated, and specified in the requirements documents. This Guide has been developed for the project officer within the Directorate of Training Developments who is responsible for preparing requirements documentation. The project officer's job responsibilities, including interfacing with other agencies with the school and elsewhere, are also described.

Handbook for Training Developers

A very recent contribution to the area is entitled: How to Determine Training Requirements and Characteristics: A Handbook for Training Developers (Lenzycki and Finley, 1980). This handbook was prepared for Army training developers who are accountable for determining training device requirements and characteristics. It describes a method for performing various front-end analyses to determine training requirements. This handbook also helps training developers determine the training device requirements, that is, whether a training device is required as a medium for developing task skills. Finally, the handbook helps to identify characteristics which a device should have to promote effective learning.

These two methods represent a logical front-end extension to such models as TRAINVICE and TRAINVICE II. They essentially provide training analysts with guidance on the essential inputs to transfer of training models.

RESULTS

During the early stages of this effort, several changes were made to the TRAINVICE II formula. One involved changing the original TRAINVICE II Index Formula to the following:

$$\text{TRAINVICE II Index} = \frac{\Sigma \left(\dfrac{PC + FC}{PC_{max} + FC_{max}} \right) \left(C \times P \times D \right)}{\Sigma \left(P \times D \right)}$$

Where:

C = Coverage Score

P = Training Proficiency Score

D = Learning Difficulty Score

PC = Physical Characteristics Score

FC = Functional Characteristics Score

PC_{max} = Maximum Possible Physical Characteristics Score

FC_{max} = Maximum Possible Functional Characteristics Score

The rationale for the changes in the formula basically involved the effect of the Coverage Analysis (C) Score. Discussion with the TRAINVICE II developers, as well as reviews of the related research revealed that the Training Criticality and Training Difficulty analyses could essentially be considered as weighting factors for required skills and knowledges. Thus, a skill or knowledge which is required to be achieved at a high level of proficiency and is also difficult to learn, should be considered to have more significance than a skill which requires a lower level of proficiency and/or is relatively easy to learn.

The model was originally conceived to accomodate this weighting factor according to the level of importance (i.e., difficulty and/or proficiency) or the skill or knowledge. For example, given two devices under consideration where one covers an important skill and the other does not, the evaluation model should penalize the device which fails to cover that skill.

Obviously, the device which adequately covers the skill or knowledge should not be penalized. The original TRAINVICE II index formula failed to accommodate this situation. In applying the original TRAINVICE II index formula, a device failing to cover a required skill (i.e., CR score = 1), would receive a zero "0", for the Coverage Analysis Score corresponding to that skill. When a training analyst calculates the index for that skill, the zero C Score would cause both the numerator and the denominator of the formula to go to zero. The result would be as if that skill or knowledge never existed. The revised TRAINVICE II index formula does penalize a device for failing to cover a required skill. In the revised formula, a C Score of zero causes only the numerator, but not the denominator, to go to zero. The denominator increases as a function of the Proficiency and Difficulty scores, with the result being a lowering of the index because of a failure to cover a required skill or knowledge. This revision of the formula accomplishes the intent of the developer as well as improving the effectiveness of a device evaluation.

Other modifications in the model included changing the names of some components. The Training Criticality Analysis (C_i) in the original TRAINVICE II model was essentially an estimate of the desired level of proficiency judged to be required of a trainee. The name of the component seemed to be misleading to a potential user, thus the term, Training Proficiency (P), was substituted.

A third major change concerned the Difficulty score. The original measurement scale in the Training Difficulty (D) Analysis was adapted from Rankin (1975). It was:

1 = Minimal or none

2 = Some

3 = Much

4 = Substantial

The revision of TRAINVICE II employs the following scale:

1 = EASY: trainee can accomplish the activity once informed that it exists; virtually no practice or study is required.

2 = MODESTLY DIFFICULT: trainee can accomplish most of the activity subsequent to instruction with little practice or study; some of the activity does require minimal practice/study to sustain competent performance at the desired level of proficiency.

3 = DIFFICULT: trainee can accomplish the activity following instruction, but only with consistent practice and/or study.

4 = HIGHLY DIFFICULT: trainee requires "extensive instruction, practice and/or study to accomplish the activity; requirements of learning at least border on expert performance standards.

This revision is thought to provide additional guidance to a user in assigning a D value to a skill or knowledge.

The Training Proficiency (P) scale also was revised. In the original TRAINVICE II model, the following scale was used to assign the P score:

1 = Should have limited knowledge of subject or skill

2 = Should have received complete briefing on subject or skill

3 = Should have understanding of subject or skill to be performed.

4 = Should have complete understanding of subject, or be highly skilled.

In TRAINVICE II, users rate each skill or knowledge in terms of the level of proficiency expected of a trainee after training. It was thought that the definitions for the rating scale were somewhat vague and may thus tend to attenuate the reliability of the model. That is, because of the lack of specificity of the definitions, judgments made by different users may vary according to their interpretations. A new scale was substituted, as shown below:

1 = Should have limited knowledge of subject or skill; has at least been briefed on the subject or performed the skill once; however, system effectiveness would probably be seriously degraded by incorrect performance at this level of proficiency.

2 = Should have minimally competent knowledge of subject or skill for performing job or operating system; some errors in performance occur regularly, but basically the individual student can sustain a minimally acceptable or "novice level" or performance.

3 = Should have adequate knowledge of subject or skill to assure reliable performance; errors in performance are infrequent to rare; performance can be characterized as smooth and experienced.

4 = Should have expert-level knowledge of subject and/or outstanding skill capability; errors in performance are rare; performance is excellent/superior.

While it is recognized that the changes made in both the D and P analysis scales do not achieve an optimum level of sophistication, they do tend to provide some additional guidance to users. For additional comments on related issues, refer to the Recommendations section of this report.

A final change to the original TRAINVICE II model involved the Behavioral Category Definitions adapted from TRADOC Pamphlet 350-30.

The TRAINVICE II model originally applied Learning Categories extracted directly from the TRADOC publications with no guidance for their use. A user was, therefore, required to make a decision regarding categories to which a skill or knowledge belonged. For example, in conducting a Physical Characteristics Analysis, a user is asked to assign a skill to a behavioral category. Under each behavioral category are listed a number of "good" instructional practices. The user is then asked to evaluate displays and controls employed in the teaching of that skill. A comparison is then made between the displays and controls, and the good instructional practices listed under the behavioral category associated with the selected skill. A potential problem was realized in attempting to assign skills or knowledges to a behavior category in that no definitions or guidance existed on how to accomplish this activity. The revision supplies definitions and examples of activities associated with each behavioral category. This additional guidance is considerably more explicit and, hopefully, provides users with a less ambiguous task; however, this area remains as a major problem with the model.

Guidebook Description

TRAINVICE II is an analytic method for assessing training devices and training device concepts. The model aids users in determining how well a training device covers and/or implements specific training requirements. TRAINVICE II can be applied either to the assessment of an existing training device or early in the development and acquisition cycle. This project was designed to develop a User's Guidebook to document the components of TRAINVICE II in a format which enables Army users to apply the model in training equipment acquisition and/or evaluation activities.

Application of TRAINVICE II to existing training devices is termed a "predictive application" of the model, and is the primary focus of the Guidebook. A "prescriptive application" of the model in the formulation of a training device concept is also possible, and is described in Appendix C of the Guidebook.

The TRAINVICE II model consists of six separate components. The components are designed so that a user can apply each separately, or can combine them into a complete TRAINVICE II analysis by computing an index to show the overall potential of a training device or simulator.

In order to aid users in applying TRAINVICE II in total or in part, each TRAINVICE II component is addressed in a separate chapter of the Guidebook. To conduct a complete TRAINVICE II analysis, a user should read all chapters in the Guidebook and following the procedures described. If, however, users are interested only in one or more of the components, they may proceed directly to the appropriate chapter and follow the procedures for that part of the TRAINVICE II model. It is recommended that users become familiar with the entire TRAINVICE II model before attempting to apply any component individually.

Each Guidebook chapter is organized in the same fashion. The first part of each chapter provides a definition and discussion of the subject component. This is followed by a presentation of the procedure used for the particular TRAINVICE II component. Where appropriate, this section discusses the relationship of the subject component to other components. This discussion is followed by a detailed example of an application of the component. Two hypothetical training devices are used as examples. All applicable skills and knowledges are listed. The user is thus able to follow an application of TRAINVICE II across the same skills and knowledges for all components, similar to an actual application. Finally, each chapter provides a summary which highlights the major points covered.

The following features have been incorporated in the Guidebook.

- Pages are numbered within chapters
- Chapters have flow charts and figures, as appropriate, to show the sequence of operations required for completing the TRAINVICE II components. The flow charts fold out for easy reference while reading the text. The flow charts show precisely where a user is in the TRAINVICE II process.

- Major points are highlighted in a box; other important points are identified by bullets to make them stand out.
- Examples are provided for illustration.

Issues of Concern

Aspects of the TRAINVICE II model represent methodological advances in the assessment of training devices for transfer of training, and the sequence of TRAINVICE II calculations appears to be logical and straightforward.

However, the calculation of the TRAINVICE II index appears to be somewhat unorthodox. The index involves a variety of types of data in what appears to be a series of simple fractions. The basic rules for adding fractions are, however, violated. For example, in calculating the PC and FC analyses, denominators are simply cumulated over skills and/or knowledges with no attempt to find a common denominator. This may be a cause for ambiguity in the model. The basis upon which a user calculates the PC_{max}, FC_{max}, PC and FC scores is different for each component. Different guidelines on "good instructional practices" are also employed for each component. The procedure for deriving the PC and FC scores is, at best, somewhat cumbersome. The consequences of this problem may well result in lower user reliability than would otherwise be the case.

Of primary concern is the cumbersomeness of the calculations for an actual device. Some training devices may address literally hundreds of skills and knowledges. It will require excessive lengths of time to derive the TRAINVICE II index in such a case, since TRAINVICE II requires a user to add, and to multiply long series of numbers, and only then to compute the final division. Arithmetic errors, therefore, may be a major factor in the model's reliability.

Concern also exists about the terminology involved in the PC and FC analyses. The immediate impression conveyed is that the Physical

Characteristics analysis has something to do with the concept of physical fidelity. Such an interpretation is, however, incorrect. The PC analysis has little or no relationship to the concept of physical fidelity, and this concept is in fact missing from TRAINVICE II. Additional components which also are missing include: psychological (functional) fidelity, criticality, learning deficit, and user acceptability.

The TRAINVICE II model requires users to generate a so-called "Consolidated List" of training skills and knowledges, which are derived from the list of all operational skills and knowledges. In essence, this consolidation eliminates the redundancy in skills and knowledges which are repeated over more than one task or subtask. Computation of the TRAINVICE II index is conducted using the Consolidated List. An argument can however, be made that if a skill or knowledge appears in more than one task it is, therefore, important by definition. A transfer of training model might be advised to somehow account for this occurrence rather than to minimize the importance of such a skill or knowledge.

Another concern involves the use of the Learning Guidelines for "Good Instructional Practice" extracted from TRADOC Pamphlet 350-30. These guidelines were developed for training programs in general, as opposed to training devices in particular. Their relationship and use in training device evaluation, particularly in transfer of training measures, is suspect. These guidelines are at best extremely difficult to apply in TRAINVICE II. They simply do not relate adequately to training device situations.

Another major issue concerns the required input to a TRAINVICE II analysis, namely task analysis. No guidance is provided in TRAINVICE II concerning the task analysis requirements for generating the necessary list of skills and knowledges. In order to make an informed Coverage Requirements Analysis (CR) decision, some information should be provided to help a user in considering this issue. As it presently stands,

no guidance in assigning the CR score is provided in TRAINVICE II. A CR score of 0, effectively eliminates a skill or knowledge from further analysis, perhaps (in some cases) erroneously.

User acceptability is another concept which is missing from TRAINVICE II. Although this concept may be difficult to measure, an otherwise excellent training device is essentially useless if it is not accepted by trainees and instructors. It appears reasonable to consider inclusion of a user acceptability measure as a weighting factor in a transfer of training model. Given two equally adequate training devices, the device found more acceptable by users will probably encourage greater transfer.

While an attempt is made to provide guidance in calculating the Training Proficiency and Difficulty scores in TRAINVICE II, these areas also remain vague. The major concern, again, involves reliability of measurement across users. The Proficiency and Difficulty components are both critical to the TRAINVICE II model. As mentioned, they function as weights for each skill and knowledge. Consistency in assigning values, therefore, is of major concern.

The final product of the TRAINVICE II analysis is the index computation. The TRAINVICE II index computation results in a figure of merit ranging between 0 and 1. When two or more devices are compared, a decimal number representing their respective merit is obtained for each device. In the predictive application of TRAINVICE II, the figures of merit for each device are compared; with the larger indicating presumably greater transfer of training. The index value, however, has no inherent meaning. It is simply an ordinal rating which is derived for each device. It is unknown what magnitude constitutes a significant difference. When two indices are very different, relative decisions among devices appear simple also; however when differences are small, relative decisions about training transfer are decidedly difficult. Further research is required to clarify the meaning of the index and to establish levels of significance for differences.

Another area of concern involves overall model reliability and validity. The concept of reliability here addresses consistency of measurement. In other words, will the same index value be computed by different users for the same device? The reliability of TRAINVICE II is simply unknown. More specific scoring techniques may be helpful for assigning component values, however, empirical research is warranted.

The question of the model's validity involves two related issues: predictive validity and construct validity. Predictive validity addresses the issue of whether the index is useful in predicting relative effectiveness of training devices. Construct validity is related to the degree to which components of the model hypothesized to influence transfer, actually measure what they are intended to measure. Construct validity addresses the structure of the model as opposed to the output. Each TRAINVICE II component, as well as the overall model must be validated. Validation of the TRAINVICE II model is a major and pressing need.

RECOMMENDATIONS

Because of increased emphasis on the development and utilization of training devices and on the rising costs associated with their development, a need exists for a systematic analytical methodology for use in assessing training devices and concepts. Such a model should also be capable of being applied in the development cycle of material systems. TRAINVICE II purports to be such a methodology. It provides a framework for the application of judgemental data concerning aspects of a training situation which are hypothesized to effect transfer of training from instructional to operational equipment. TRAINVICE II has several potential areas of application. These include:

- Identification of potentially unsatisfactory devices before they are built.
- Identification of acceptable changes or modifications proposed for training devices.
- A selection methodology for use in choosing among competing devices.

The purpose of this project was to review the existing TRAINVICE II documentation and to convert the model into a user-oriented Guidebook format. TRAINVICE II, in its present form represents an advance in training device assessment. There are, however, a considerable number of problem areas with the model in its present state.

Before additional effort is expended in further revision of the TRAINVICE II model, several important steps should occur. Because of the cumbersomeness of calculating the final index, an automated version of TRAINVICE II should be developed. Training devices very often cover many, perhaps hundreds, of skills and knowledges. A user, attempting to apply the present manually calculated index, will find the procedure extremely time consuming and tedious. The ready availability of microprocessors and their low cost make them natural candidates for such an

application. Models currently available, such as the TRS-80, are sufficiently fast, and appear to provide for adequate storage facility. In addition, the TRS-80 is sufficiently portable to be made readily accessible to users. Computerizing TRAINVICE II on a micro-processor is recommended.

A second recommendation which could be accomplished in parallel with the first, involves a complete revision of the guidelines used in the PC and FC analyses. Such an effort would result in improved standards against which training device components (displays and controls) can be evaluated. Such standards should be device-specific, as opposed to general instructional guidelines, which is presently the case.

Third, further research is also required on the model components themselves. This is particularly true of the Physical and Functional Characteristics analyses.

Fourth, consideration should be given to the possible addition of other components to the model. Such topic areas as: consequences of inadequate performance, task criticality, physical and functional fidelity of simulation, user acceptance, cost, crew or team applications, etc., are sample topic areas which are currently not covered in the existing model. Consideration should be devoted to their inclusion.

Fifth, efforts should be made to empirically assess the respective contributions of each model component to the effectiveness of the overall model in both its predictive and prescriptive applications.

Sixth, research should be devoted to the scale properties of the index figure of merit.

Seventh, each model component should be assessed empirically for reliability and validity, and the entire model should undergo an empirical validity study prior to its implementation.

REFERENCES

Adams, J., On the Evaluation of Training Devices, Human Factors, 1979, 21(6), 711-720.

Kern, R., Sticht, T., Welty, D., and Hauke, R., Guidebook for the Development of Army Training Literature, U.S. Army Research Institute Technical Report, 1977, AD AO 33935

Klein, G. C., Kane, J. J., Chinn, A. A., and Jukes, A. O.. Analyzing Training Device Effectiveness In Cases Where Test Data is Inconclusive. Proceedings of The Human Factors Society - 22nd Annual Meeting, 1978

Lenzycki, H. P., and Finley, D. L. How To Determine Training Requirements and Characteristics: A Handbook for Training Developers. U. S. Army Research Institute for The Behavioral and Social Sciences. Technical Report (draft) 1980.

Narva, M. A. Formative Utilization of a Model for the Prediction of the Effectiveness of Training Devices. U. S. Army Institute for the Behavioral and Social Sciences, Research Memorandum 79-6, May 1979a.

Narva, M. A. Development of a Systematic Methodology for The Application of Judgemental Data To The Assessment of Training Device Concepts. U.S. Army Research Institute for the Behavioral and Social Sciences, Research Memorandum 79-7. May 1979b.

Rankin, W. Task Description and Analysis for Training System Design. U. S. Navy, Training Analysis and Evaluation Group, Orlando, Florida. Technical Memorandum 74-2, 1975.

Swezey, R. W., Chitwood, T. E., Easley, D. L. and Waite, B. J. Implications for TOW Gunnery Training Development, Springfield, Virginia. Mellonics Systems Development Division, Litton Systems, Inc. 1977a (2 volumes).

Swezey, R. W., Chitwood, T. E., Easley, D. L., and Waite, B. J. Implications for Dragon Gunnery Training Development, Springfield, Virginia, Mellonics Systems Development Division, Litton Systems, Inc. 1977b (2 volumes).

Swezey, R. W., and Evans, R. A. Guidebook for Users Of TRAINVICE II. McLean, Virginia, Behavioral Sciences Research Center, Science Applications, Inc., 1980.

Swezey, R. A., and Narva, M. A. (EDS.) Proceedings: Conference on The TRAINVICE Model for Assessing Training Device Effectiveness, Alexandria, Virginia, Allen Corporation of America, 1978.

Training Device Requirements Documents Guide. A Procedures Handbook for Directorate of Training Development Project Offices for Devices. PM Trade, Naval Training Equipment Center, 1979.

U.S. Naval Training Device Center. Staff study on cost and training effectiveness of proposed training systems. TAEG Report 1, U.S. Naval Training Center, Orlando, Florida, 1972.

Wheaton, G. R., Rose, A. M., Fingerman, P. W., Korotkin, A. L., Holding, D. H., and Mirabella, A. Evaluation of the Effectiveness of Training Devices: Literature Review and Preliminary Model. US Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA, Research Memorandum 76-6, 1976a.

Wheaton, G. R., Fingerman, P. W., Rose, A. M., and Leonard, R. L. Evaluation of the Effectiveness of Training Devices: Elaboration and Application of the Predictive Model. US Army Research Institute for the Behavioral and Social Sciences. Research Memorandum 76-16, 1976b.

Wheaton, G. R., Rose, A. M., Fingerman, P. W., Leonard, R. L. and Boycan, G. G. Evaluation of the Effectiveness of Training Devices: Validation of the Predictive Model US Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginia, 1976c.